

특징선 정합을 이용한 3차원 얼굴 모델링

김항기^U 김황수
경북대학교 컴퓨터과학과
hangkee@hanmail.net, hsk@knu.ac.kr

3D Face Modeling Using Feature Line Fitting

Hang-Kee Kim^U Hwang-Su Kim
Dept. of Computer Science, Kyungpook University

요 약

본 논문에서는 3차원 머리 모델에 몇 장의 사진으로부터 얻은 텍스처를 입혀 실물처럼 보이는 3차원 인물 모델을 얻는 방법을 제시한다. 모델에 사진들을 맞추는 방법으로는 특징선을 정합하는 방법을 사용한다. 모델에는 얼굴의 특징을 나타낼 수 있는 눈/코/입/눈썹 등의 특징선을 지정하였으며 이들을 사진에 정합시킴으로써 모델의 각 부위에 필요한 텍스처 영상을 얻는다. 여러 방향에서 본 사진들을 사용함으로써 더욱 정확한 얼굴 모델을 얻을 수 있는데, 이때 모델의 한 면은 여러 장의 사진에서 합성되어야 하는 경우가 생긴다. 이는 각 사진에서 얼굴이 보는 방향과 모델면이 이루는 각을 이용하여 그 사진이 그 면의 텍스처에 기여하는 정도를 계산할 수 있다. 이렇게 함으로써 사진을 이용한 저가의 3차원 캡처 시스템을 구현할 수 있다.

1. 서론

3차원 얼굴 애니메이션은 가상현실, MPEG-4영상압축, 화상회의, 영화, 광고, 엔터테인먼트 분야 등에서 많이 활용되고 있다. 특히 하드웨어, 소프트웨어의 발달에 따라 실시간으로 3차원 캐릭터를 움직일 수 있을 정도까지 되었다.

현재 상용화되어 있는 3차원 모션캡처 장비는 사실성과 정확성 면에서 아주 우수하지만, 매우 고가인데다가 일반인들이 사용하기 어려운 단점이 있다. 따라서 화상회의 시스템이나, 가상 에이전트에서의 3차원 캐릭터 구성이 아직은 힘든 시점에 있다.

그러나 CCD같은 저가형 비디오 캡처 시스템은 일반 가정에서도 쉽게 사용할 수 있고, 그에 따라 비디오 캡처를 통한 화상회의, 애니메이션, 얼굴인식 등이 가능하게 되었다.

따라서 본 연구에서는, 저가형 비디오 캡처 시스템을 이용한 사실적인 3차원 애니메이션을 구현하는 제반 시스템을 개발하는 것에 목표를 두고자 한다. 또한 정면, 측면뿐만 아니라 비디오 캡처를 통한 모든 방향의 인물 영상을 이용하여 실제 인물을 더 정확하게 3차원 모델로 구성할 수 있게 한다.

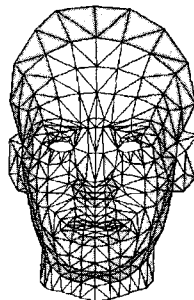
관련된 연구로는 70년대 초 Ekman[1]이 얼굴에 대한 감정, 표정을 각각 기본적인 Action Unit 단위로 구분하여 그 조합으로 표현한 FACS(Facial Action Coding System) 방법론을 제시하였고, Parke[2]가 컴퓨터를 이용한 얼굴 애니메이션 방법을 발표한 이후, 90년대에 이르러서 Waters[3]는 근육을 기반으로 한 3차원 얼굴

애니메이션을 만들었으며, Thalmann, Essa등에 의해 다양한 방법으로 연구되고 있다.

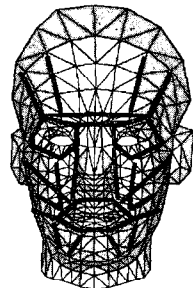
3. 와이어프레임 모델링

3.1. 근육 기반 모델링

사실적이면서도 자연스러운 애니메이션을 위해서는 얼굴의 해부학적 지식을 바탕으로 한 근육기반 모델링이 필요하다[3]. 본 논문에서는 근육기반 모델링을 위한 와이어프레임 모델로써 <그림1>과 같은 모델을 이용한다. 이 모델은 381개의 꼭지점과, 730개의 삼각면, 36개의 근육이 사용되었다.



(a) 정면



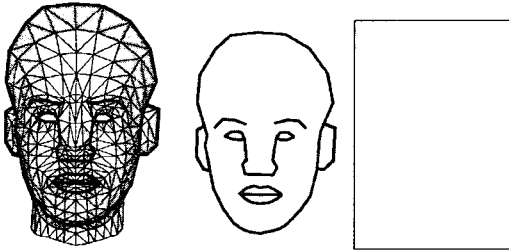
(b) 모델에 적용된 근육

<그림 1> 본 논문에서 제안한 와이어프레임

3.2. 특징선 제어

사람의 얼굴에서 식별 가능한 특징점에는 눈, 코, 입, 눈썹, 얼굴 외곽선 등이 있다. 2차원 사진에서 이와 같은 특징점을 이용하여 얼굴을 정합시키는 것이 편리하다. 따라서 <그림2>와 같은 특징선을 모델에 지정한다.

이들 특징선을 얼굴에 정합시킬 때 수작업에 의한 방법과 특징점 추출에 의한 자동화된 방법이 있다. 현재는 특징선을 사용자가 직접 이동시키는 수작업에 의한 다.



(a)모델+특징선 (b)특징선 (c)사진+특징선
<그림2>모델에 적용된 특징선

3.2.1.사용자의 입력에 의한 특징선 제어

사용자의 입력에 의하여 특징선을 제어할 경우 보다 정확하고 세밀한 제어를 할 수 있으나 사용자의 번거움을 피할 수 없다. 따라서 모델의 모든 점을 제어하기 보다는 특징선만 얼굴에 정합시키고 나머지 점들은 특징선을 기준선으로 한 Warping으로 구현할 수 있다. 본 논문에서는 사용자와의 대화적 특성을 위하여 속도가 빠르다고 할 수 있는 Inverse Distance Weighted Function을 사용하였다[4]. 이 방식을 간략히 설명하면, 다음과 같다.

내부의 각 점 p 와 특징선이 되는 점들 p_i 와의 상대적인 거리인 가중치 w_i 를 다음 식을 이용하여 구한다.

$$w_i(p) = \frac{1}{\|p - p_i\|^\mu}$$

여기서 지수 μ 는 거리의 영향을 조절하는 파라미터이다. 이 가중치는 x 와 y 좌표를 연속적으로 보간하는 데 사용된다.

$$W_x(p) = \frac{\sum_{i=1}^N w_i(p) p_x}{\sum_{i=1}^N w_i(p)}$$

최종 warp된 위치는 $W(p) = (W_x(p), W_y(p))$ 이 된다.

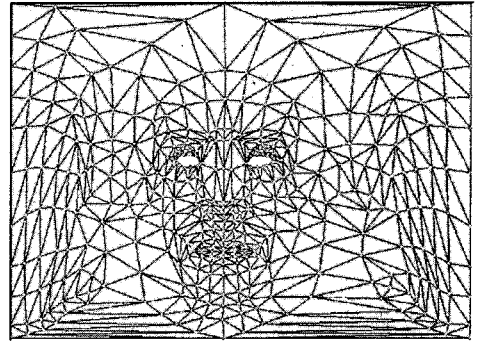
3.2.1.특징점 추출에 의한 특징선 제어

지금까지 얼굴 인식 논문에 나온 여러 가지 얼굴 특징점 추출방식을 이용하면 눈, 코, 입, 얼굴 영역에 대한 특징들을 찾을 수 있다[5]. 찾아진 특징들에 주어진 모델의 특징선을 정합시킬 수 있다.

4. 텍스처 맵핑

4.1. 한 사진으로부터 텍스처 영상 얻기

한 사진과 와이어프레임 모델이 특징선 제어를 통하여 정합되고 나면 와이어프레임 모델 중 보이는 면들의 텍스처를 그 사진으로부터 얻을 수 있다. 보이는 면의 삼각형들을 텍스처 맵 영상에 Warping기법을 이용하여 구한다. 사진에서 가져온 각 면의 텍스처가 전체 텍스처 맵에 맵핑될 위치는 <그림 3>과 같다.



<그림 3> Texture Map 영상의 좌표

4.2. 텍스처 결합

특징선 제어를 통하여 정합된 와이어프레임 모델을 이용하여 보이는 면에 대한 텍스처를 얻을 수 있다. 이때 사진을 하나의 3차원 평면이라고 봤을 때, 그 3차원 평면의 법선 n 과 얼굴 모델의 각 면의 법선 u 와의 각도 차이를 이용하여 그 사진이 주어진 면에 얼마만큼 영향을 끼치는지를 계산할 수 있다. 즉, n 과 u 의 내적의 비율로써 텍스처 맵에 대한 기여도를 계산할 수 있다. 따라서 모델의 어떤 면 F_j 에 대하여 N 개의 사진중 각 사진의 텍스처 $t_i(F_j)$ 를 이용하여 면 F_j 의 텍스처인 $T(F_j)$ 를 구할 수 있다. 이때 μ 는 내적의 차이가 미치는 영향을 조절하기 위하여 사용된다.

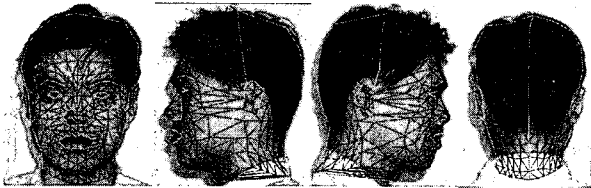
$$T(F_j) = \sum_{i=1}^N \frac{(n_i \cdot u_j)^\mu}{\sum_{k=1}^N (n_k \cdot u_j)^\mu} t_i(F_j)$$

위의 방법으로 각 면에 대한 Texture를 구하면 각 면들 간에 합성 비율의 차이로 인하여 경계 단절 현상이 일어난다. 따라서 각 면에 대한 비율을 계산할 것이 아니라, 텍스처 맵 영상의 각 픽셀단위로 비율을 계산해 주어야 한다. 텍스처 맵 영상의 각 픽셀에 대한 노멀을 구하려면 먼저 Phong Shading에서 노멀을 계산하는 방식과 마찬가지로 각 면의 꼭지점에 대한 노멀을 주위 면들의 노멀의 평균을 이용하여 구한다. 그 후 그 픽셀의 노멀을 꼭지점과의 관계를 이용하여 구한다.

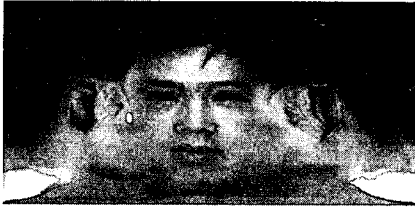
5. 실험 및 결과 분석

5.1. 전/후/좌/우 4방향 영상을 이용한 모델링

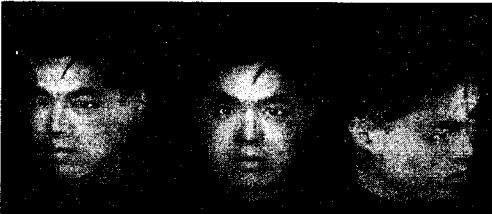
제안한 방법은 Pentium III CPU에, OpenGL가속보드가 설치된 시스템에서 테스트를 하였으며, 테스트 영상은 디지털 카메라를 이용하여 <그림4>와 같은 전/후/좌/우 4방향의 영상을 얻은 다음 특징선 정합으로 모델링 하였다. 특징선 제어는 사용자가 모델의 특징점들을 이동시켜줌으로써 정합되었고, 특징선 외의 점은 Warping에 의해서 계산된다. 텍스처 영상은 자동으로 4방향에서 얻은 텍스처를 합쳐 <그림5>와 같은 텍스처 맵 영상을 만들어 준다. <그림6>에서 볼 수 있듯이 실제 사진과 매우 닮은 모델을 얻을 수 있을 뿐만 아니라 그 외의 방향에서 보았을 때에도 자연스러운 모습을 얻을 수 있다.



<그림4> 모델링에 사용된 얼굴의 디지털 카메라 사진



<그림5> 그림4의 사진을 이용하여 만든 Texture Map



<그림6> 3차원 얼굴로 재구성된 장면

5.2. 자유 방향 영상을 이용한 모델링

두 번째 테스트 방법은 CCD 카메라를 이용하여 자유로운 방향에서 찍은 동영상을 이용한 모델링 방법이다. 그 예는 <그림 7>과 같다. <그림 8>에서 볼 수 있듯이 전/후/좌/우 4방향을 이용하여 모델링하였을 때보다 부자연스러운 모양이 나타남을 알 수 있다. 이는 CCD 카메라 자체의 왜곡현상이 있는 관계로 한쪽 방향에 정합시킨 후 다른 방향서 맞추려면 모양이 맞지 않기 때문에 와이 어프레이밍 정합이 곤란해져서이다. 이 문제는 카메라 calibration을 통하여 해결할 수 있을 것으로 보인다.



<그림7> CCD 카메라로 찍은 얼굴사진



<그림8> 위 CCD 사진이 3차원으로 재구성된 장면

6. 결론

본 논문에서는 2차원 영상에 특징선을 정합시킴으로써 3차원 머리 모델링을 하는 방법을 제시하였다. 이를 이용하면 고가의 3차원 스캐너가 없이도 자연스러운 머리 모델링이 가능하다. 또한 여러 방향에서 찍은 사진의 Texture를 합쳐 Texture Map 영상을 만들 때 모델의 각 면에서 가장 크게 영향을 미치는 사진의 Texture에 비중을 더 주게 하여 더 자연스럽게 세밀한 영상이 나올 수 있게 하였다. 따라서 이를 저가형 화상회의, video-mail, animation 등에 활용할 수 있을 것이다. 향후 연구과제로서는 특징점을 자동으로 추출하여 얼굴 정합을 스스로 하는 방법을 연구하고, 나아가서 실시간 얼굴 추적을 통한 애니메이션을 구현하고자 한다.

참고 문헌

- [1] P. Ekman, W.V. Friesen, "Facial Action Coding System". Consulting Psychologists Press Inc., 577 College Avenue, Palo Alto, California 94306, 1978
- [2] Frederic I. Parke, Computer Generated animation of faces, MS thesis, Univ. of Utah, Salt Lake City, UT, December 1974.
- [3] Yuencheng Lee, Demetri Terzopoulos, Keith Waters, 1995, "Realistic modeling for facial animation", SIGGRAPH95, p55-62
- [4] Shepard, D. 1968. A Two-Dimensional Interpolation Function for Irregularly Spaced Data. Proc. 23rd National Conference of the ACM, 517-524
- [5] Takacs, B., Wechsler, H., 1998. A Saccadic Vision System for Landmark Detection and Face Recognition. Face Recognition, NATO ASI Series VOL 163, p627-636